



## Cancer Research Center Hotline

### Laboratory Analysis of HER-2/*neu*: Search for the Optimal Testing Strategy

**Jeffrey L. Killeen MD**

**Director of Laboratories**

**Kapiolani Medical Center for Women and Children  
Pan Pacific Pathologists, Inc.**

Over the last decade increases in the understanding of human carcinogenesis have been remarkable. While this knowledge is far from complete, the identification of specific genetic and molecular events promises to radically alter the way cancers are treated by elucidating molecular targets for new pharmaceutical agents. A few such drugs have already been developed including Gleevec, targeting c-kit; Iressa, targeting epidermal growth factor receptor; and Herceptin, targeting the HER-2/*neu* (HER-2) protein. These agents show promise as primary treatment or as adjuvant therapy combined with standard treatments such as surgery, radiation therapy, chemotherapy and hormonal therapy. This new experience with "pharmacogenomics" has also brought new challenges for the diagnostic laboratory. The proper identification of patients who may benefit from these new drugs is critical. Equally important is identifying who is not a good candidate as these drugs may have serious side effects and are very expensive. This new developing pathology practice of identifying and characterizing therapeutic targets has been termed "pharmacopathology".<sup>1</sup> The current controversy surrounding the laboratory evaluation of HER-2 status is the prime example of this new dilemma in diagnostic pathology. It is worthwhile for pathologists and clinicians alike to examine this debate for it is a preview of what may well follow as more molecular based therapies advance through the clinical trial phase and into clinical practice. The HER-2 debate is an object lesson in why close interaction between the clinician and the pathologist is becoming more and more important. It is imperative that the pathologist becomes more familiar with the management implications of diagnostic tests. Likewise, clinicians must realize that there is more to pathologic evaluation than a hard copy result. Proper interpretation of pathologic results, whether they be from your local lab or data appearing in the peer-reviewed literature, requires an understanding of the testing process. The HER-2 story is a complex one that involves not only laboratory practice but also significant marketing and economical concerns. These "non-medical" forces must be understood as they clearly affect what appears in the peer-reviewed literature on this subject and directly shape marketing campaigns to clinicians and pathologists. In this commentary I will attempt to summarize the information, and misinformation, currently available and propose a rational approach to HER-2 testing.

The current debate focuses on what is the preferred method for HER-2 evaluation. While several methods have been employed since the discovery of the HER-2 gene and protein in 1985<sup>2,4</sup>, the two clinically relevant methods are immunohistochemistry (IHC) and

fluorescence in situ hybridization (FISH). IHC is used to measure receptor protein expression and FISH is used to measure HER-2 gene copy number. The controversy centers around the reliability of IHC-based HER-2 testing and whether FISH should be the testing method of choice. To sort out this conundrum it is best to first review a few aspects of HER-2 biology, the laboratory issues in the development of anti-HER-2 therapy and the processes of immunohistochemistry and fluorescence in situ hybridization.

Overexpression of the HER-2/*neu* gene can be demonstrated in 20-30% of invasive breast cancers.<sup>5,6</sup> The gene is located on chromosome 17q and encodes a transmembrane tyrosine kinase receptor related to epidermal growth factor. In most, but not all, cases overexpression of the HER-2 protein (called p185) is associated with amplification of the HER-2 gene. The importance of HER-2 alterations is two-fold in that it has both prognostic and predictive value. Prognostically, tumors with overexpression or amplification are associated with shorter disease-free and overall survival while retrospective studies have shown that HER-2 alteration may be predictive of response to certain types of chemotherapy.<sup>7</sup> Most importantly, the HER-2 protein is the target of trastuzumab (Herceptin), the humanized monoclonal antibody now available for the treatment of breast cancer. Optimal use of this drug requires accurate determination of HER-2 status.

In 1993, as a result of clinical data from the Herceptin trial, it was apparent that a commercially available IHC test was necessary to determine patient selection for Herceptin. At that time, practical gene-based tests were not available. During the trial a combination of two antibodies called the "Clinical Trials Assay (CTA)" had been used to detect overexpression. The IHC staining was semi-quantitatively assessed on a 4-tiered scale (0, 1+, 2+, 3+) by evaluating the amount and intensity of cell membrane staining. Pre-clinical work suggested that Herceptin activity required a receptor density greater than 100,000 to 200,000 per cell. This roughly corresponded to an IHC score of at least 2+ by CTA so patients with 2+ or 3+ scores were considered candidates for Herceptin while those with scores of 0 or 1+ were considered non-overexpressed. The CTA assay was costly and difficult to perform making it an impractical test commercially. Genentech, the makers of Herceptin, granted Dako Corporation a license to develop an IHC kit that could be commercially used to detect HER-2 overexpression. The original tumor samples from the patients in the Herceptin trial had deteriorated so the newly developed antibody and testing protocol (HercepTest) along with the CTA were applied to a large set of tumors obtained from the National Cancer Institute. The tumor set was specifically chosen to have equal numbers of positive (2+, 3+) and negative (0, 1+) cases. This is a far more rigorous scenario than actual clinical practice where approximately 80% of cases will be clearly normal (0, 1+) or clearly overexpressed (3+). The comparison study showed a concordance between CTA and HercepTest of 79%. Based on this, the FDA unanimously approved HercepTest as the method for determining patient eligibility for Herceptin. The HercepTest procedure is more than just the polyclonal antibody (clone A0485). In an attempt to minimize variation from lab to lab the kit also includes all reagents along with a specific testing protocol and detailed interpretive guide.

Because most of the criticism of the IHC method of HER2 assessment is centered around the reported variability of the IHC method, a brief review of the IHC process is necessary. The IHC

method for identifying specific proteins in histologic preparations has been common in the pathology laboratory for over 20 years and is performed daily in most labs. The method consists of multiple steps, all of which can be modified by the laboratory to optimize the desired result. These processes include tissue fixation, antigen enhancement techniques, incubation times and conditions, dilution of the primary antibodies and detection methods (linking molecules and chromogens). In essentially all antibodies routinely used by pathologists the primary goal is to determine if a certain protein is present or not. For example, IHC stains for keratin may be used to determine if a tumor displays evidence of epithelial origin. The pathologist does not care *how much* keratin is present, just whether it is there or not. Because of this "yea or nay" approach procedures for specific IHC stains vary from lab to lab, depending on lab preference. Antibodies themselves vary in that different "clones" are available which may detect similar, but not identical, epitopes. No less than seven antibodies directed against various epitopes of the HER-2 protein are commercially available.<sup>8,9</sup> It is not surprising that interlaboratory variability results in this setting but with "qualitative" IHC this is not a major problem. The IHC determination of HER-2, however, is unlike any other IHC test. Because HER-2 is also present on normal breast epithelial cells it is necessary to quantitate the number of receptors present and establish a cut-off point above which "over-expression" is present. This concept of "quantitative" IHC is new for pathologists. While the HercepTest kit introduced the concept of quantitative IHC, it was also far more expensive for laboratories to utilize compared to other HER-2 antibodies. As a result, many laboratories altered the procedures, both in testing conditions and in the antibodies used. Many of these deviations formed the basis of HER-2 results reported in the literature and it is not surprising that variation resulted.

With this background on the development and variation in the laboratory methods of IHC we can look at the criticisms of this method of HER-2 assessment. A major factor adding to the intensity of the discussion was the availability of an alternative testing method, fluorescence in situ hybridization (FISH). Because of the stability of DNA, FISH testing could be performed on tumor samples previously tested by IHC and the results compared. In spite of what is implied in FISH marketing tactics, the procedure is not simple. It includes its own set of variables including tissue fixation, testing conditions, probe characteristics and interpretive criteria. In addition, FISH is significantly more expensive and performed by far fewer labs than IHC. This latter point is the primary reason that variability studies done to date have suggested FISH is more reproducible than IHC. As FISH testing becomes more widespread, interlaboratory variation will become more of a problem. When the IHC/FISH comparison studies were done, concerns quickly arose regarding the accuracy of IHC, and by implication, the reliability in patient selection for Herceptin therapy. This latter issue, that of patient selection, is especially important for Genentech, the manufacturer of Herceptin. If the test used to select the patients most likely to respond is flawed the perceived efficacy of the drug is at risk. As a result, Genentech has been active in the questioning of the IHC method and has been promoting FISH methodology.

The most commonly cited study by those advocating FISH over IHC for HER-2 testing is the report by Pauletti et al.<sup>10</sup> In that study 900 cases of invasive breast cancer diagnosed between 1987 and

1991 in South Australia were tested for HER-2 status by IHC and FISH with results compared with patient survival. Both methods were independent predictors of survival on multivariate analysis. By univariate analysis, there was a direct correlation between survival and gene copy number determined by FISH while that same relationship was only seen in the strongest staining IHC group (3+). This study has been put forth as evidence that FISH is the more accurate test. Keeping in mind the intricacies of IHC methodology discussed earlier, it is worthwhile taking a closer look at the Pauletti study. This re-analysis was well described by Yaziji and Gown and a few pertinent findings stated here.<sup>11</sup> The authors of the Pauletti study used the R60 polyclonal antibody which is not widely available and has been infrequently studied. They did not use antigen retrieval which is considered standard in most labs today. They utilized the peroxidase-antiperoxidase detection system which was common in the 80's and early 90's but has since been replaced by more sensitive methods such as avidin-biotin or streptavidin-biotin. These technical problems with the Pauletti study likely introduced both false positive and false negative IHC results leading to the suboptimal performance of their IHC assay. Another study raising concerns about the IHC method is the National Surgical Adjuvant Breast and Bowel Project (NSABP) Protocol B-31 which looked at the benefit of adding Herceptin to chemotherapy.<sup>12</sup> Eligibility was based on HER-2 results from the accruing institutions and included cases that were either 3+ by HercepTest, showed strong membrane staining if other HER-2 antibodies were used, or gene amplification by FISH. Central HER-2 retesting was performed for the first 104 cases entered and showed that 18% of the community-based assays could not be confirmed by central HercepTest or FISH assay. The authors questioned the reliability of IHC performed in lower volume laboratories since repeat IHC and FISH testing at the central, large volume lab showed concordance in 94% of cases.

A third example of problems with IHC can be found in the experience of the Breast Intergroup Trial N9831 which also is evaluating Herceptin in the adjuvant setting.<sup>13</sup> A secondary endpoint of the trial is to assess concordance between HercepTest and FISH at a central testing facility. Eligibility was the same as the NSABP B-31 trial. Of the first 119 patients entered, 110 had local IHC testing while only nine had local FISH assays. On central IHC repeat testing only 81 of the 110 (73%) showed 3+ staining with HercepTest. This again raised serious concerns about the accuracy of IHC, particularly when performed in lower volume community laboratories. What is not emphasized by those who cite this study as evidence of FISH superiority is that local vs. central FISH assays also showed poor concordance. Of the nine patients entered with local FISH testing showing gene amplification, only six (67%) were confirmed on central FISH testing. When both IHC and FISH were performed by the central lab there was 92% concordance. The 8% discordance consisted of cases that were 3+ by HercepTest but non-amplified by FISH, a consistent subgroup that will be mentioned later.

The publication of these studies and some others led to questioning of the reliability of IHC for determining HER-2 status. This effort has been supported by Genentech and those with interests in FISH technology. Subsequently many published reports have followed showing that IHC is, in fact, very reliable when strict adherence to testing and interpretive protocols are followed.<sup>14,15</sup> It has also been clear for some time that the "2+" IHC category should

not be considered "over-expressed" and requires additional testing. In spite of this fact, the proponents of FISH continue to cite studies that included 2+ reactions as "positive" as evidence of the shortcomings of IHC. An example of the excellent results that can be obtained when high quality IHC is performed can be seen in the data presented by Yaziji and Gown.<sup>11</sup> Among the 381 tumors negative for amplification by FISH, only six (1.6%) showed 2+ staining on IHC. Of the cases positive for overexpression on IHC (3+), 94% were amplified by FISH. This and many other studies published in the last 3-5 years as well as data from our own lab confirms that IHC, when well performed is highly accurate in determining HER-2 status.

As you can see from this discussion the large majority of the literature addressing the issue of IHC vs. FISH for HER-2 testing consists of comparison studies and retrospective studies looking at HER-2 as a prognostic factor. As mentioned previously, many of these studies are seriously flawed by variation in IHC testing practice and inconsistency in how results are reported. What has been lacking until recently is data showing the direct relationship between HER-2 testing methods and patient response to Herceptin. This important data is now starting to appear in the literature. In data submitted to the FDA from the Herceptin Clinical Trial (H0648g) significant improvement in time to progression was seen in patients receiving Heceptin plus chemotherapy compared with those receiving chemotherapy alone. This benefit was not only seen in patients who were IHC 3+/FISH+ but also in a subgroup of patients who were IHC 3+/FISH-, giving a false negative FISH rate of 13%.<sup>16</sup> In a study evaluating Herceptin as a single first-line agent in patients with metastatic breast cancer, 29 patients who were IHC 3+ responded to Herceptin. Twenty-seven patients with FISH positive tumors responded however 2 patients who were FISH negative (7%) also showed objective response.<sup>17</sup> Most recently, in a phase II monotherapy trial from Germany, clinical response to Herceptin was limited to those patients with IHC 3+ tumors. Most of these were also FISH positive but three patients who responded were FISH negative, a FISH false negative rate of 15%.<sup>18</sup> Clearly, the most recent and clinically relevant data shows that a strongly positive IHC test correlates best with patient response. This only makes sense given that the target of Herceptin is the p185 surface protein, the protein measured by IHC quantitative testing. In this sense FISH should be looked at as a "surrogate" test for Herceptin eligibility. It appears there is a small (7-15%) but consistent group of patients who are IHC+/FISH- that will respond to Herceptin. This may be due to chromosome 17 polysomy or low-level amplification with FISH HER-2 ratios between 1.0 and 2.0.

It seems then that when we look for the optimal testing method for HER-2 status we have come full circle. We began with IHC and now it is clear that IHC, indeed, is the best first-line test. What we have learned during the journey is that quality laboratory practice is more critical than the method itself. A laboratory that performs IHC in a sloppy fashion will likely perform FISH just as poorly. If a laboratory cannot ensure that the appropriate procedures are followed and that HER-2 results are continually monitored and data analyzed then samples should be referred to labs that have these procedures in place. The College of American Pathologists is currently structuring a HER-2 quality control program to help labs address this issue. While FISH is an alternative to IHC in first-line testing, a more appropriate role is in confirming borderline or "2+" IHC cases. This

approach utilizing IHC as the initial test followed by reflex FISH testing for indeterminate cases has now been supported by many groups and is included in the recent National Comprehensive Cancer Network (NCCN) Guidelines.<sup>16,19</sup>

In our laboratory we have optimized the interpretation of HercepTest by applying digital computerized image analysis. The ACIS image analysis system (ChromaVision Medical Systems, San Juan Capistrano, CA) utilizes custom software to analyze HercepTest IHC stains. This technique is more accurate than manual scoring and improves reproducibility.<sup>20,21</sup> Automated IHC analysis will likely become more common as the area of pharmacopathology expands.<sup>22</sup> In conclusion, the saga of HER-2/neu testing is complex but necessary to understand for it is a vision of the future where multidisciplinary interaction, critical literature review, and bit of healthy skepticism are needed to determine the path to optimal patient care. While HER-2 testing will continue to evolve, current data indicates the optimal testing strategy should employ quality IHC with reflex FISH testing of indeterminate results.

For information on the Cancer Research Center of Hawaii, please visit [www.crch.org](http://www.crch.org).

## References

1. Schnitt SJ. Pharmacopathology: A New Spin on an Old Idea. *Am J Surg Pathol*, 2003;27:121-123.
2. Coussens L, Yang-Feng TL, Liao YC, et al. Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. *Science* 1985;230:1132-1139.
3. Akiyama T, Sudo C, Ogawara H, et al. The product of the human *c-erbB-2* gene: A 185-kilodalton glycoprotein with tyrosine kinase activity. *Science* 1986;232:1644-1646.
4. Stern DF, Heffernan PA, Weinberg RA. P185, a product of the *neu* proto-oncogene, is a receptorlike protein associated with tyrosin kinase activity. *Mol Cell Biol* 1986;6:1729-1740.
5. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/*neu* oncogene. *Science* 1987;235:177-182.
6. Slamon DJ, Godolphin W, Jones LA, et al. Studies of HER-2/*neu* proto-oncogene in human breast and ovarian cancer. *Science* 1989;244:707-712.
7. Ravdin PM, Charnness GC. The *c-erbB-2* proto-oncogene as a prognostic and predictive marker in breast cancer: A paradigm for the development of other macromolecular markers - A review. *Gene* 1995;159:19-27.
8. Bilous M, Dowsett M, Hanna W, et al. Current Perspectives on HER2 Testing: A Review of National Testing Guidelines. *Mod Pathol* 2003;16:173-182.
9. Rhodes A, Jasani B, Anderson E, et al. Evaluation of HER-2/*neu* Immunohistochemical Assay Sensitivity and Scoring on Formalin-Fixed and Paraffin-Processed Cell Lines and Breast Tumors. *Am J Clin Pathol* 2002;118:408-417.
10. Pauletti G, Dandekar S, Rong H, et al. Assessment of Methods for Tissue-Based Detection of the HER-2/*neu* Alteration in Human Breast Cancer: A Direct Comparison of Fluorescence In Situ Hybridization and Immunohistochemistry. *J Clin Oncol* 2000;18:3651-3664.
11. Yaziji H, Gown A. Testing for HER-2/*neu* in Breast Cancer: Is Fluorescence In Situ Hybridization Superior in Predicting Outcome? *Advances in Anatomic Pathology* 2002;9:338-344.
12. Paik S, Bryant J, Tan-Chiu E, et al. Real-World Performance of HER2 Testing - National Surgical Adjuvant Breast and Bowel Project Experience. *J Natl Cancer Inst* 2002;94:852-854.
13. Roche PC, Suman VJ, Jenkins RB, et al. Concordance Between Local and Central Laboratory HER2 Testing in the Breast Intergroup Trial N9831. *J Natl Cancer Inst* 2002;94:855-857.
14. Kakar S, Puangsuvan N, Stevens JM, et al. HER-2/*neu* Assessment in Breast Cancer by Immunohistochemistry and Fluorescence In Situ Hybridization: Comparison of Results and Correlation With Survival. *Molecular Diagnosis* 2000;5:199-207.
15. Seidman AD, Fornier MN, Esteva FJ, et al. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol* 2001;19:2587-2595.
16. Official Transcripts. FDA ODAC Meeting. December 5, 2001.
17. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and Safety of Trastuzumab as a Single Agent in First-Line Treatment of HER2-Overexpressing Metastatic Breast Cancer. *J Clin Oncol* 2002;20:719-726.
18. Hofmann MK, Gross C, Beyser K, et al. Predictive value of HercepTest IHC and Pathvysion FISH data: Analysis from trastuzumab phase II monotherapy study. *Proc. ASCO* 2003;22:1185.
19. National Comprehensive Cancer Network (NCCN) Practice Guidelines, Version 1, 2003.
20. Wang S, Saboorian MH, Frenkel EP, et al. Assessment of HER-2/*neu* Status in Breast Cancer: Automated Cellular Imaging System (ACIS)-Assisted Quantitation of Immunohistochemical Assay Achieves High Accuracy in Comparison With Fluorescence In Situ Hybridization Assay as the Standard. *Am J Clin Pathol* 2001;116:495-503.
21. Bloom KJ, Assad L. Validating HER-2/*neu* assessment by immunohistochemistry. Abstract 239, 25<sup>th</sup> Annual San Antonio Breast Cancer Symposium. December 11-14, 2002.
22. Seidat T, Balaton AJ, Battifora H. Interpretation and Quantitation of Immunostains. *Am J Surg Pathol* 2001;25:1204-1207.